

Non-line-of-sight tracking of people at long range

SUSAN CHAN,^{1,2} RYAN E. Warburton,^{1,2} GENEVIEVE GARIEPY,¹
JONATHAN LEACH,¹ AND DANIELE FACCIO^{1,*}

¹*Institute of Photonics and Quantum Sciences, Heriot-Watt University, David Brewster Building, Edinburgh, EH14 4AS, UK*

²*These authors contributed equally.*

*Corresponding author: d.faccio@hw.ac.uk

Abstract: A remote-sensing system that can determine the position of hidden objects has applications in many critical real-life scenarios, such as search and rescue missions and safe autonomous driving. Previous work has shown the ability to range and image objects hidden from the direct line of sight, employing advanced optical imaging technologies aimed at small objects at short range. In this work we demonstrate a long-range tracking system based on single laser illumination and single-pixel single-photon detection. This enables us to track one or more people hidden from view at a stand-off distance of over 50 m. These results pave the way towards next generation LiDAR systems that will reconstruct not only the direct-view scene but also the main elements hidden behind walls or corners.

Published by The Optical Society under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

OCIS codes: (150.5670) Machine vision, Range finding; (040.1345) Optoelectronics, Avalanche photodiodes (APDs).

References and links

1. A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, "Looking around the corner using transient imaging," in *Proceedings of IEEE International Conference on Computer Vision* (IEEE, 2009), pp. 159–166.
2. R. Pandharkar, A. Velten, A. Bardagjy, E. Lawson, M. Bawendi, and R. Raskar, "Estimating motion and size of moving non-line-of-sight objects in cluttered environments," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2011), pp. 265–272.
3. A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nature Comms* **3**, 745–748 (2012).
4. O. Gupta, T. Willwacher, A. Velten, A. Veeraraghavan, and R. Raskar, "Reconstruction of hidden 3D shapes using diffuse reflections," *Opt. Express* **20**, 19096–19108 (2012).
5. F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, "Low-budget transient imaging using photonic mixer devices," *ACM Trans. Graph.* **32**, 45–50 (2013).
6. M. Buttafava, J. Zeman, A. Tosi, K. Elceiri, and A. Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," *Opt. Express* **23**, 20997–21115 (2015).
7. M. Laurenzis, J. Klein, E. Bacher, and N. Metzger, "Multiple-return single-photon counting of light in flight and sensing of non-line-of-sight objects at shortwave infrared wavelengths," *Opt. Lett.* **40**, 4815–4818 (2015).
8. G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio, "Detection and tracking of moving objects hidden from view," *Nature Photon.* **10**, 23–26 (2015).
9. S. Shrestha, F. Heide, W. Heidrich, and G. Wetzstein, "Computational imaging with multi-camera time-of-flight systems," *ACM Trans. Graph.* **35**, 1–11 (2016).
10. A. Kadambi, H. Zhao, B. Shi, and R. Raskar, "Occluded imaging with time-of-flight sensors," *ACM Trans. Graph.* **35**, 1–12 (2016).
11. J. Klein, C. Peters, J. Martín, and M. Laurenzis, "Tracking objects outside the line of sight using 2D intensity images," *Scientific Reports* **6**, 32491 (2016).
12. U. Wandering, "Introduction to Lidar," in *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere*, C. Weitkamp, ed. (Springer, 2005), pp. 1–18.
13. C. Niclass, A. Rochas, P.-A. Besse, and E. Charbon, "Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes," *IEEE J. Solid-State Circuits* **40**, 1847–1854 (2005).
14. F. Zappa, S. Tisa, A. Tosi, and S. Cova, "Principles and features of single-photon avalanche diode arrays," *Sens. Actuators A Phys.* **140**, 103–112 (2007).
15. J. Richardson, R. Walker, L. Grant, D. Stoppa, F. Borghetti, E. Charbon, M. Gersbach, and R. K. Henderson, "A 32x32 50ps resolution 10 bit time to digital converter array in 130nm CMOS for time correlated imaging," in

- Proceedings of IEEE Custom Integrated Circuits Conference* (IEEE, 2009), pp. 77–80.
16. E. Charbon, M. Fishburn, R. Walker, R. K. Henderson, and C. Niclass, “SPAD-Based Sensors,” in “TOF Range-Imaging Cameras,” F. Remondino and D. Stoppa, eds. (Springer, 2013), pp. 11–38.
 17. J. A. Richardson, L. A. Grant, and R. K. Henderson, “Low dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology,” *IEEE Photon. Technol. Lett.* **21**, 1020–1022 (2009).
 18. G. Gariepy, N. Krstajić, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio, “Single-photon sensitive light-in-flight imaging,” *Nature Comms* **6**, 6021–6026 (2015).
 19. W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques* (Springer, 2005).
 20. R. Cutler and L. Davis, “View-based detection and analysis of periodic motion,” in *Proceedings of International Conference on Pattern Recognition* (IEEE Comput. Soc, 1998), pp. 495–500.
 21. R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, “Detecting moving objects, ghosts, and shadows in video streams,” *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 1337–1342 (2003).
 22. T. S. Ralston, G. L. Charvat, and J. E. Peabody, “Real-time through-wall imaging using an ultrawideband multiple-input multiple-output (MIMO) phased array radar system,” in *Proceedings of IEEE International Symposium on Phased Array Systems and Technology* (IEEE, 2010), pp. 551–558.

1. Introduction

Recent advances in light sensing and computational imaging technologies are providing solutions to the problem of looking around obstacles [1–11]. In particular, devices that can detect the arrival of light at the single-photon level with extremely high temporal resolution have enabled the ranging and reconstruction of images of small, static hidden objects using methods based on laser-illuminated detection and ranging (LiDAR) [3, 6, 8]. LiDAR is an active remote-sensing technique that uses the round-trip time-of-flight information of light signals backscattered from objects, typically in the line of sight, to determine their positions [12]. LiDAR-based detection and ranging of objects hidden from view is achieved by illuminating the objects and detecting the backscattered signals via an intermediary scattering surface such as a wall or the floor. These additional, intermediate scattering events and their isotropic nature greatly reduce the available signal for detection, leading to the need for long acquisition or processing times, and/or the need for advanced detection devices.

Buttafava *et al.* recently studied the possibility of determining the full three-dimensional profile of static objects using a single-pixel single-photon avalanche diode (SPAD) and scanning laser illumination [6]. Although the system did not have the necessary speed to track a moving object in real-time, single-pixel SPADs do have a distinct advantage over SPAD cameras [13–16] in that they provide close to 100% coupling of light onto the sensitive detector area, compared to the few percent currently available in visible–near-infrared wavelength SPAD cameras [17, 18].

For some applications, object identification and position or motion information are the main points of interest; a three-dimensional reconstruction of the object is not crucial [2, 8]. Rather, the scenario may demand, first and foremost, knowledge of the presence and position of objects moving in a hidden environment. Example applications include search and rescue, and autonomous driving. Gariepy *et al.* have shown that it is possible to detect and track a moving hidden object, albeit with no information of the object’s form [8]. Their setup uses a 1024-pixel SPAD camera [13, 15, 17] to detect indirect light scattered back to the system from their target object around a corner. Each detector pixel in the 32×32 array has single-photon sensitivity and can time each photon’s arrival with 110 ps resolution. Tracking of the hidden object uses this precise photon-arrival timing information together with knowledge of the camera’s field of view and the geometry of the setup in the line of sight.

Here we present a tracking system that extends beyond the lab-based near-range setups implemented thus far. Our system enables the detection of people moving outside of the direct line of sight at stand-off distances greater than 50 m. The active imaging system uses one single-pixel SPAD detector and a single pulsed laser to “look” down a corridor and around a blind corner. In a first experiment, we demonstrate that we can accurately detect and locate a single person around the corner. We then demonstrate that we can accurately do this for two hidden people.

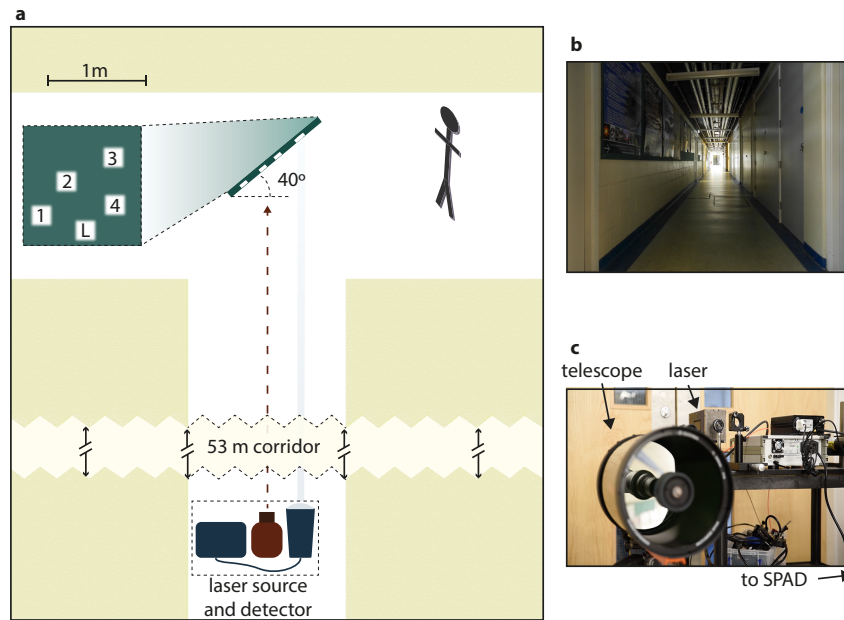


Fig. 1. Looking down the corridor. We perform our experiment by looking down the corridor at a screen. (a) shows a schematic overview of the geometry of the environment. The blind corners obstruct direct vision of the hidden person(s). In a single measurement, we illuminate a single spot on the wall with a high-repetition pulsed laser and detect the light scattered back to the points on the wall which we image to our SPAD detector. “L” indicates the position of the laser on the screen and the numbered squares indicate positions imaged (one at a time) onto the single-pixel SPAD. (b) shows a photograph of the environment. Lighting conditions are as in the experiment. (c) shows a photograph of the equipment.

The expected decrease in spatial resolution due to the use of a single- or few-pixel detection system compared to previous camera-based approaches is compensated by the arbitrarily large effective numerical aperture that is provided by using independently focusable detectors: even moving objects at large distances can be located with high precision. These results, therefore, show that single-photon counting technology can indeed perform outside the lab and over length scales that are relevant for large stand-off distance detection and location of hidden moving objects, such as humans, in real-life applications.

2. Experimental setup

In our experiments we use a corridor that forms a T-junction. Figure 1 shows a schematic of the scenario as seen from above. Compared to the experiment conducted by Garipey *et al.* where an intermediary scattering surface in the plane of the object’s movement (the floor) was used, here we use a vertical surface (screen), as in previous work (see e.g. [1–4]). Our screen is a mobile chalkboard that we can freely position at the T-junction. The laser and the field of view of the detector are directed onto viewing screens attached to the board to improve the signal level of our returns.

The transceiver comprises a pulsed laser diode (Picoquant LDH-P-FA-1530, 1532 nm peak wavelength, 40 MHz repetition frequency, ~ 400 mW average power), a time-correlated single-photon counting [19] (TCSPC) module (Picoquant PicoHarp 300), an InGaAs/InP SPAD (ID Quantique ID230) and a commercial telescope (Celestron Schmidt-Cassegrain, 8-inch aperture,

80-inch focal length). The returning light is coupled to the SPAD through a $62.5\ \mu\text{m}$ diameter core fibre. A 1500 nm long-pass filter was used. The timing resolution of the detector is ~ 200 ps.

We send a train of light pulses from the transceiver onto our screen positioned at the far end of the corridor (~ 53 m away from the transceiver), as illustrated in Fig. 1(a). The 100 ps pulses from the portable laser hit the wall with a ~ 7 cm beam diameter and scatter from the wall approximately as a spherical wavefront that propagates in all directions. Some of this light reaches the hidden objects (persons) and is scattered back again towards the wall. A discrete position with a spot size diameter of ~ 4 cm on the wall is imaged to the SPAD using the telescope. Our TCSPC module measures the photon arrival time (4 ps time binning) for the signal returning to the detector and a histogram is built up in one second of acquisition time over 40 million laser pulses.

3. Position retrieval

In order to precisely locate a hidden object, we successively image four discrete positions on the screen to the SPAD as indicated in Fig. 1(a). This is equivalent to simultaneously imaging four detector positions to four SPADs with their associated collection optics; i.e. with four independent detectors and imaging systems, we would, in effect, be able to perform the total data acquisition in 1 s. We use the temporal information between the laser signal and the SPAD detector signal for each of the four pixels (detector positions) to reconstruct the position of the hidden persons. In our system, we can simply choose our point of reference (origin of the

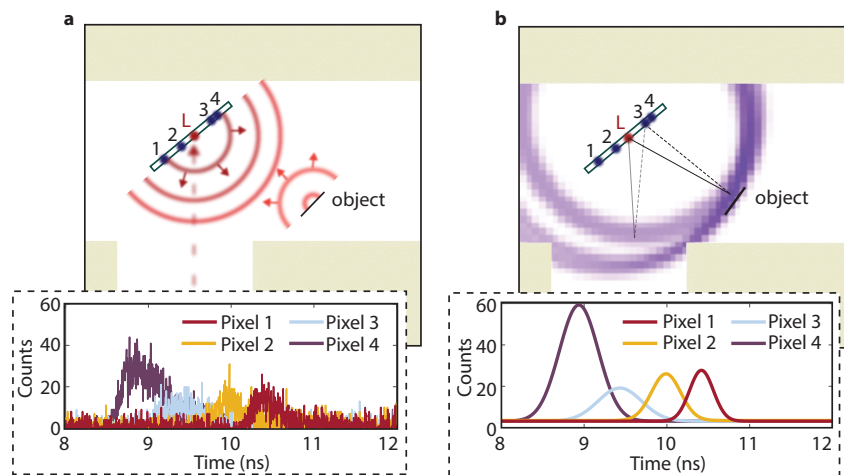


Fig. 2. Target-position retrieval. (a) Laser pulses incident on the wall scatter and propagate approximately as an isotropic spherical wavefront. Light reaching an object scatters back in a similar fashion. The inset histogram shows an example of the signal recorded by the single-pixel detector. (b) The time extracted from the peaks (of the Gaussian fits to the data, shown in the inset) provide the total distance travelled by the light, from the wall to the person and back to the wall, but not the actual path taken. The bold and dashed lines in the schematic show two equivalent paths, corresponding to the time extracted from the histogram for one pixel position - there are infinite such paths that combined, form an ellipsoid of equally probable locations for the hidden object. By repeating the measurement four times (looking at different positions on the screen as shown in Fig. 1(a)) we obtain four ellipses that overlap in correspondence to the target position in two dimensions.

Cartesian coordinate system), for example the righthand-side corner of the T-junction, and we then use an extension of the approach presented in [8].

To calibrate our measurements, we apply a global offset to the histograms recorded for each pixel (detector position) and define the timing of all events relative to the start of the histogram. We then limit the histogram to our time window of interest. In our current method, we pre-acquire a signal of the background scene with the person(s) absent, for each detector pixel, and subtract this from the return signal for the corresponding pixel when we have the person(s) present in the scene. Alternatively we can consider using the median of the histograms for each pixel [8, 20, 21] or frame-to-frame change detection [2, 22] as a way to approximate the background signal.

For each detector pixel i , we perform a Gaussian fit on the histogram to locate the position of the peak(s) in the return signal corresponding to our scattering source(s) of interest (see insets in Fig. 2). Each fitted peak gives us our photon time-of-flight t_i , with uncertainty σ_i , that we measure between the initial instant when the train of laser pulses hits the wall at r_l and the final instant when it has scattered to and from a hidden person at r_o , back into the field of view of the detector at r_i . Using t_i , our method back-projects the photons onto the hidden scene. We create an ensemble of discrete positions in Cartesian space for where our hidden person could be located. Defining an appropriate scattering height z for our person reduces the search space from a volume (x, y, z) to an area (x, y) with a significant increase in computation speed. For each discrete position $r_o = (x_o, y_o, z)$, we then calculate a probability:

$$P_i(r_o) \propto \exp \left[-\frac{(|r_o - r_l| + |r_o - r_i| - ct_i)^2}{2\sigma_i^2} \right] \quad (1)$$

The probability density that we obtain for our two-dimensional problem is maximised for a set of (x, y) that describes an ellipse with $r_l = (x_l, y_l, z_l)$ and $r_i = (x_i, y_i, z_i)$ as the foci, such that $|r_o - r_l| + |r_o - r_i| = ct_i$. Here, c is the speed of light. Assuming independence between the measured times of flight t_i , the final distribution of r_o is obtained by multiplying the probability density functions (PDFs) $P_i(r_o)$ associated with each pixel. Since the counts in our target signal have contributions from the same scattering source, the resulting distribution will present a region of high probability where the individual densities overlap.

4. Results

4.1. Tracking of a single person

A single person is standing around the righthand-side corner, hidden from the line of sight of our transceiver. The screen is angled at ~ 40 degrees to the corridor on the right to mimic, for example, an open door. The corridor lights are switched off yet full daylight illumination enters from lateral windows, as seen in Fig. 1(b), however no significant differences were observed with the artificial corridor lights switched on or off. We acquire data for 1 s with the person located at each of four positions along the corridor, and we do this for all four pixels. The signal of the background scene is also acquired for each pixel. Figure 3 shows the joint probability density functions (PDF) that we retrieve for the hidden person at each position. These are overlaid with the ground truth. The agreement between the PDFs and the person's true positions show that our system is able to locate a stationary person situated up to ~ 1.8 m outside of the direct line of sight, and determine its position with an uncertainty that is always less than ~ 0.5 m, although this increases the further away the person is from the screen.

In order to assess how location accuracy and precision varies with the baseline (i.e. the separation of pixel positions on the screen), we perform a computer simulation of the experiment for the simplified case where we have a single laser, two SPADs and a hidden object (see Fig. 4). For four arbitrary object locations, we scan the pixel position of one of our detectors (D_2) in the x -direction to see how this affects the retrieved position. The results indicate that a baseline of 1 m

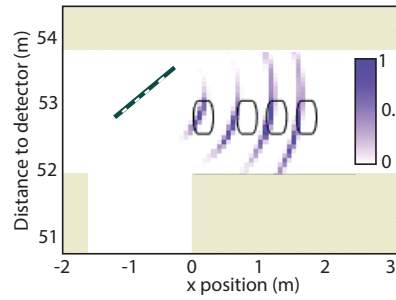


Fig. 3. Experimental results for non-line-of-sight tracking of a single person. We perform tracking of a hidden person at distinct positions around a blind corner. Each coloured curve in the graph is a joint probability distribution of the person's retrieved position, while the corresponding rectangle is their actual position during the measurement.

is sufficient to provide an average location accuracy of ~ 0.1 m in the x -direction and ~ 0.25 m in the y -direction (see Figs. 5(a) and 5(b)), and an average location precision of ~ 0.4 m and ~ 0.7 m in the x - and y -directions respectively (see Figs. 5(c) and 5(d)). By introducing further pixels, a 1 m baseline will therefore be sufficient to guarantee a location precision comparable to the typical shoulder width of a human being. In most cases, the precision increases very rapidly with increasing baseline but then tails off. In other words, there is not much gain in going to large baselines - baselines of the order of 1 m are optimal in terms of requirements on the environment (access to a sufficiently large visible area) and location accuracy. These results are verified in our experiments in which we have baselines of the order of 1 m.

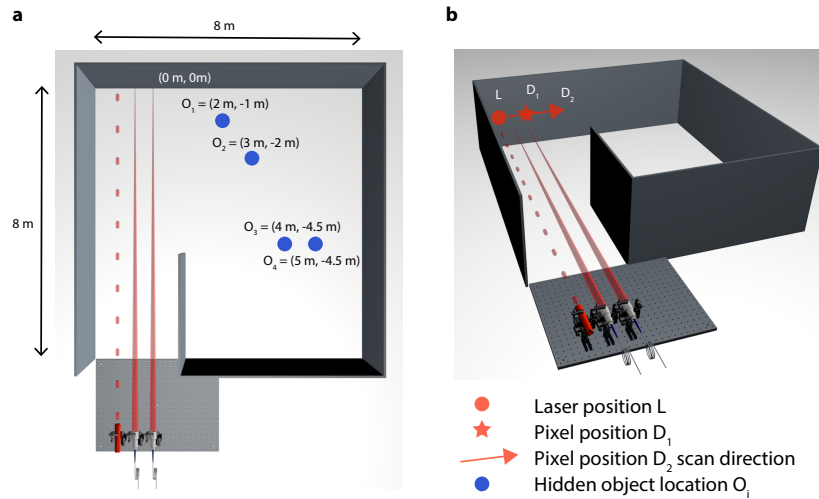


Fig. 4. Simulation setup for a two-detector system. To investigate the baseline distance between the detector positions on the screen, we simulate an experiment with a single laser and two detectors. (a) and (b) show a schematic overview of the simulated environment in top-down and perspective view respectively. Laser position "L" and pixel position "D₁" are fixed while pixel position "D₂" is scanned in the x -direction. " O_i " indicates the location of the hidden object in each set of simulated measurements.

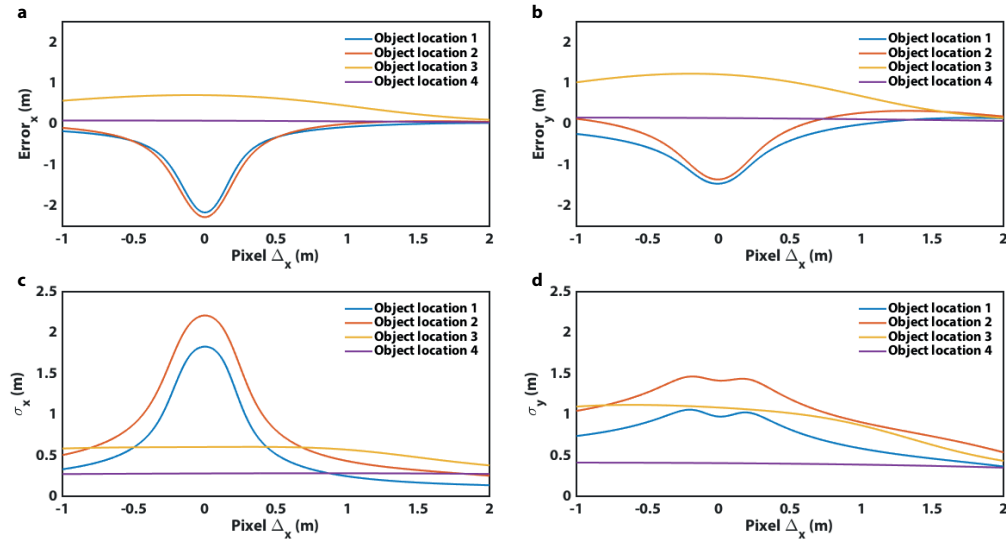


Fig. 5. Simulation results for a two-detector system. We investigate how separating out the pixels in one direction affects the accuracy and precision of the target position retrieval. (a) and (b) show the respective change in accuracy of the retrieved x - and y -coordinates ("error _{x} " and "error _{y} ") as the separation between pixels increases, while (c) and (d) show the change in precision in the x - and y -directions (" σ_x " and " σ_y ") respectively.

4.2. Tracking of multiple persons

The previous experimental measurements are repeated with two people in the same hidden scene. Again, data is acquired for 1 s for each detector position. We show the results in Fig. 6. Figures 6(a) and 6(b) show examples of the actual data and corresponding Gaussian fits used for the retrieval, highlighting the presence and also clear difference between the return signals from each person. The retrieved positions, compared to the ground truth positions, are shown in Figs. 6(c) and 6(d) respectively. As before, the positions of the people are always located to within ~ 0.5 m for distances up to ~ 2 m from the screen.

We also perform a measurement with two people located on opposite sides of the T-junction. We use only two detector positions and the screen is now placed at normal incidence with respect to the transceiver. The retrieved positions are shown in Fig. 7. The system is now retrieving the target positions with larger uncertainty and we observe that if either of the persons moved to further than ~ 2 m away from the laser spot on the wall, the signal-to-noise ratio of the return signal degraded to the point that position reconstruction is no longer possible within the integration time. Preliminary measurements showed that person orientation has no notable effect on scattering. Instead, we attribute this largely to a non-uniform scattering distribution of the photons from the screen surface with a significantly lower distribution parallel to the wall itself (i.e. in the direction of the hidden persons in this configuration). However, the presence of hidden people and their correct positions can still be identified by the system, showing that no particular tuning of the angle of the screen is required. That is, the limitations observed here can, in principle, be overcome by using detectors with lower noise (the detector used here has a dark count rate of roughly 1000 counts/second and systems with a factor $\sim 10\times$ less are available). The signal may also be increased by collecting the return scatter from a larger spot size, e.g. by optimising the collection optic numerical aperture. The spot size diameter could be increased by a factor $\sim 2 - 3\times$ (therefore increasing signal by a factor $\sim 10\times$) with respect to the current configuration without significantly compromising precision. This would imply (when

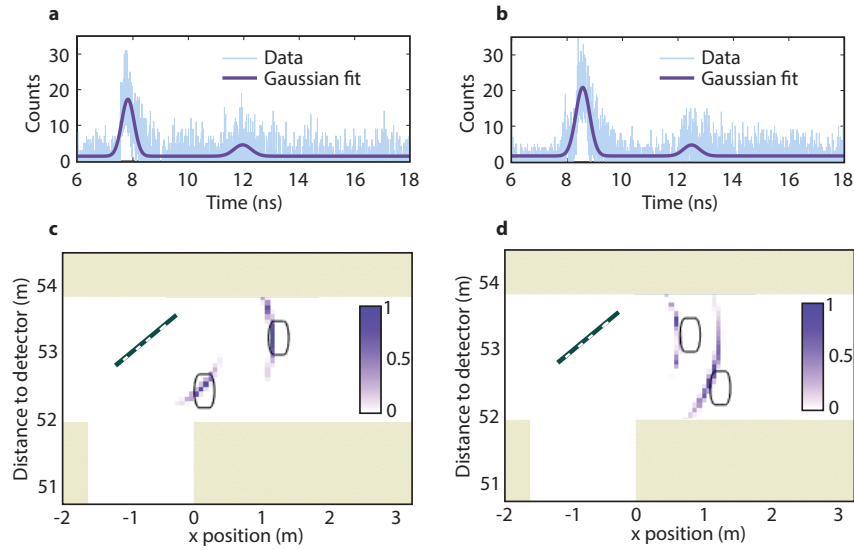


Fig. 6. Experimental results for tracking of multiple hidden persons. We perform tracking of two people located around the same blind corner. (a) and (b) show the raw data corresponding to just one of the four pixel positions used to identify and locate the two people, as shown in (c) and (d).

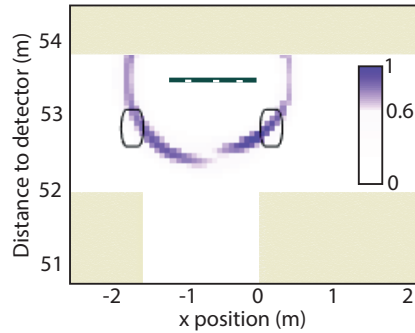


Fig. 7. Experimental results for detecting multiple hidden persons in less favourable conditions. We perform detection of two hidden persons, each located around either corner of the T-junction. The coloured curves and corresponding rectangles show each person's retrieved position and actual position relative to the righthand-side corner.

combined with lower noise detectors) a total increase in the signal-to-noise ratio of $\sim 100\times$ and the potential to detect hidden objects that are located up to several meters behind the corner while maintaining the same 50 m stand-off distance. The detected signal scales as $1/d^4$, where d is the distance between the screen and the object(s) of interest. With our current setup with the current technology, our method is limited to resolving two simultaneous objects in the hidden scene; any increase will require increased temporal resolution.

5. Discussion

Photons returning via objects located outside of the line of sight undergo three or more scattering events, yet the timing information of return signals can be used to accurately retrieve the positions

of these objects. The flexibility of single-pixel detection provides an increased field of view that allows detection and tracking with precision at long range, > 50 m. The use of single-pixel detectors also has the advantage of high detection efficiency. The results from our measurements show that by using a few single-pixel SPADs in parallel and TCSPC, real-time tracking at large stand-off distances is possible. The ability to perform non-line-of-sight detection at long range, using experimental components that can be made both compact and portable, takes us one step closer to developing a solution that is usable for real-life scenarios. An interesting possibility also is to use just one single pixel: while this will not provide sufficient information to locate the exact position of a hidden object, it is sufficient to identify the existence of a moving object, its distance from the laser beam spot, its direction of motion and velocity.

Funding

This work was funded by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC GA 306559, by the Engineering and Physical Sciences Research Council (EPSRC, UK; grants EP/M006514/1, EP/M01326X/1), and by the Defence Science and Technology Laboratory's Centre for Defence Enterprise.

Acknowledgments

We would like to thank Alessandro Boccolini for the photographs used in this paper. The authors acknowledge discussions with Ken McEwan. All experimental data is available at <http://dx.doi.org/10.17861/8a0669d3-8f9b-4d8a-b204-f980f5d57c29>.